# Supporting Information

## *Supplementary Methods*

### Ortholog detection

Orthologous groups among the EC II genomes were determined using the whole-genome alignments generated by Mauve (1), with the criteria that putative orthologs have at least 80% sequence similarity and an alignment that covers at least 45% of the larger gene. While technically the orthologs extracted from the Mauve alignments are actually positional homologs, the fact that these genomes are highly syntenic in orthologous regions makes this an effective way to identify one-to-one orthologs. The low threshold for alignment coverage was used to facilitate the identification of truncated genes caused by either mutation, sequencing error or erroneous gene prediction. For the purposes of determining the enrichment of COG categories among orthologous groups not present in all EC II genomes, suspect orthologous groups were filtered out. Specifically, orthologous groups were removed from consideration if: 1) any member was less than 90% of the length of the largest member of the group, 2) more than 10% of the sequence of any member was ambiguous bases, 3) any member of the group was within 50 bp of the end of a contig, 4) any member had fewer than 50 amino acids.

### Filtering of nucleotide differences

SNPs were filtered to remove those SNPs that are likely to be due to alignment or sequencing errors. SNPs were filtered out if: 1) they resided in genes annotated as phage, transposase or integrase, 2) they resided in genomic regions annotated as

phage by the Phage Finder program (2), 3) they resided within 20 bp of the start or end of a contig, 4) they resided in tandem repeats of total length greater than 20 bp, as determined by the exact-tandem program associated with MUMmer (3), 5) they resided in large inexact repeats as determined by nucmer, 6) they were within two positions of a second putative SNP, 6) the SNP position was ambiguous or low quality in any of the aligned genomes, 7) the 10 bp window surrounding the putative SNP contained more than two ambiguous or low quality base calls, or 8) the 10 bp window surrounding the putative SNP contained a A/T homopolymer run of length five or longer. SNPs that were informative about B and C sharing a common ancestor were verified manually by examining the underlying reads.

**Delineation of recombinant regions**

A random walk algorithm was employed to detect genomic regions of unusually high SNP density. A genome was first divided into SNP, non-SNP and indel positions, based upon a whole-genome alignment. A value, $S_i$, was then assigned to each position *i* in the genome as follows:

$$S_i = \begin{cases} \frac{1}{N_S} & SNP\ positions \\ -\frac{1}{N_N} & Non\text{-}SNP\ positions \\ 0 & Indel\ positions \end{cases}$$

where $N_S$ is the total number of SNP positions and $N_N$ is the total number of non-SNP positions. The $S$ vector was then used to calculate a context score, $C_i$, for each position *i* as follows:

$$C_i = \max\left(\left(C_{i-1} + S_i\right),\ 0\right)$$

The vector of context scores were then converted to a binary vector *R*, which indicated whether each position was in a region of high SNP density. *R* was created by first scanning through *C* to delineate the start and end of regions of high SNP density, and then updating R as follows:

$$R_i = \begin{cases} 1 & R^S_j \le i \le R^E_j, \textit{ for any } j \\ 0 & \textit{Otherwise} \end{cases}$$

where $R^S_j$ indicates the start of region *j* of high SNP density and $R^E_j$ indicates the end of region *j*. The beginnings of regions of high SNP density are assigned when scanning through the *C* vector, a value greater than a user-defined threshold *S* is observed, with this threshold position being $R^T_j$ for region *j*. $R^S_j$ is then assigned as the index of the zero value in *C* which most closely precedes $R^T_j$. The end of region *j* of high SNP density is assigned when C falls more than some user-defined threshold *E* below the maximum value observed after $R^S_j$. $R^E_j$ is then assigned the index in *C* containing this maximum. By selecting the maximum value as the end of the recombinant region, as opposed to the point where C falls below the maximum, we avoid including the trailing region of low SNP density. The search for the *j+1th* region of high SNP density begins at $R^E_j$. For demarcating recombinant regions among EC II genomes, values of *S* and *E* of 0.005 were used.

***Supplementary Results***

***Multi-locus Sequence Typing***

Multi-locus sequence typing (MLST) was performed by sequencing the regions documented in reference 35 in the main text (See main text Materials and Methods). Looking at the sequences revealed that while the outbreak strains A, B and C were identical across three of the MLST regions, in two of the regions they varied by between two and nine bases (Table S1). Sequencing the complete genomes of A, B and C revealed that the cause of this variability in the degree of divergence in different regions was that the more divergent MLST regions resided in recombinant parts of the genomes. The fact that the gdhB and gpi genes reside in recombination hotspots has been previously observed (4). It is of note that the MLST regions used in main text reference 15 do not reside in recombinant parts of the genome, and as would be expected are nearly identical among EC II genomes at all seven loci, except for a single base change in the recA gene of C (Table S1).

***Overall variation in gene content among the EC II genomes***

Not included in the main text was a comparison of the gene content in the EC II genomes (A, B, C and ACICU). For this analysis orthologs were identified using Mauve whole genome alignments. Trends in what types of genes are being gained and lost since the divergence of the EC II strains was determined by looking for enrichment in genes not present in all four genomes, across different COG functional categories. Fig. S7 shows that five categories were significantly enriched in variable genes: Replication, Recombination and Repair, Transcription, Defense Mechanisms, Cell Wall/Membrane/Envelope Biogenesis and Function Unknown. The basis for the observed variation in these classes of genes is discussed below.

*Replication, recombination and repair*

Replication, recombination and repair genes have been previously found to be variable among closely related *Acinetobacter* genomes (5), with this category encompassing transposable elements (TE's), in addition to many phage genes. This was demonstrated to be the case here, as when TE and phage genes were removed from the analysis, this functional category was no longer significantly enriched among variable genes. While transposable elements, such as insertion sequences, have been shown to be important to both the acquisition and expression of genes important to pathogenesis (REF), the draft nature of the currently sequenced genomes made it difficult to assess full impact of variable TE insertion. Phage may also carry genes relevant to pathogenesis, however, it is difficult from genome sequence alone to determine whether phage associated genes are likely to be expressed in a manner that would impact cellular function. It is clear that as has been observed in previously sequenced *A. baumannii* genomes (5-7), there was a significant phage presence. In fact, there were even differences in prophage content between the two most closely related genomes, A and HC64. It should also be noted that the enrichment for genes annotated as "Function Unknown" among variable genes could also be attributed to the large number of poorly characterized genes in prophage regions.

*Transcription*

The enrichment in transcriptional genes was in large part due to a large deletion of an ~85 Kb region in B. This large deletion may be indicative of a reduced dependence of the genes in this region, as the EC II lineage becomes more tightly associated with the hospital and host environments. Evidence for lessening importance of genes in this region came from an observed variable loss in this region among previously sequenced *A. baumannii.* However, it is of note that the only other genome in which this region was found to be completely absent is the louse associated strain, SDF.

*Defense Mechanisms*

The acquisition and loss of defense genes has been well documented in *A. baumannii* (8), making enrichment in this functional category somewhat unsurprising. Of note was observed variability in the previously described resistance islands in the EC II lineage. Specifically, we confirm an overall lesser importance of the resistance island to the MDR phenotype in EC II strains (6, 7), by observing that the primary resistance island is completely missing from A. Strain AB307-0294 of the EC I lineage also lacks this resistance island, but unlike here, loses much of its drug resistance (5). Thus it seems that this resistance island can be confer multi-drug resistance in *A. baumannii*, but is not necessary.

*Cell Wall/Membrane/Envelope Biogenesis*

The observed variability in cell wall biosynthesis genes was almost entirely a consequence of turnover of LPS biosynthetic enzymes by way of homologous recombination, as discussed in the main text.

***NISC Comparative Sequence Program Full Author List***

Jim Mullikin, Robert Blakesley, Alice Young, Grace Chu, Colleen Ramsahoye, Sean Lovett, Joel Han, Richelle Legaspi, Christina Sison, Michael Gregory, Casandra Montemayor, Marie Gestole, April Hargrove, Taccara Johnson, Jerlil Myrick, Nancy Riebow, Brian Schmidt, Betsy Novotny, Jyoti Gupta , Betty Benjamin, Shelise Brooks, Holly Coleman, Shi-ling Ho, Karen Schandler, Lauren Smith, Mal Stantripop, Quino Maduro , Gerry Bouffard, Mila Dekhtyar, Xiaobin Guan, Cathy Masiello, Baishali Maskeri, Jenny McDowell, Morgan Park, Pamela Thomas

## *Supplementary References*


1. Darling AC, Mau B, Blattner FR, Perna NT (2004) Mauve: Multiple Alignment of Conserved Genomic Sequence With Rearrangements. *Genome Research* 14:1394-1403.

2. Fouts DE Phage_Finder: Automated identification and classification of prophage regions in complete bacterial genome sequences. *Nucleic Acids Research* 34:5839 - 5851.

3. Delcher AL, Phillippy A, Carlton J, Salzberg SL (2002) Fast algorithms for large-scale genome alignment and comparison. *Nucl. Acids Res.* 30:2478-2483.

4. Hamouda A, Evans BA, Towner KJ, Amyes SGB (2010) Characterization of Epidemiologically Unrelated Acinetobacter baumannii Isolates from Four Continents by Use of Multilocus Sequence Typing, Pulsed-Field Gel Electrophoresis, and Sequence-Based Typing of blaOXA-51-like Genes. *J. Clin. Microbiol.* 48:2476-2483.

5. Adams MD, Chan ER, Molyneaux ND, Bonomo RA (2010) Genomewide Analysis of Divergence of Antibiotic Resistance Determinants in Closely Related Isolates of Acinetobacter baumannii. *Antimicrob. Agents Chemother.* 54:3569-3577.

6. Iacono M et al. (2008) Whole-Genome Pyrosequencing of an Epidemic Multidrug-Resistant Acinetobacter baumannii Strain Belonging to the European Clone II Group.

*Antimicrob. Agents Chemother.* 52:2616-2625.

7. Adams MD et al. (2008) Comparative Genome Sequence Analysis of Multidrug-Resistant Acinetobacter baumannii. *J. Bacteriol.* 190:8053-8064.

8. Peleg AY, Seifert H, Paterson DL (2008) Acinetobacter baumannii: Emergence of a Successful Pathogen. *Clin. Microbiol. Rev.* 21:538-582.

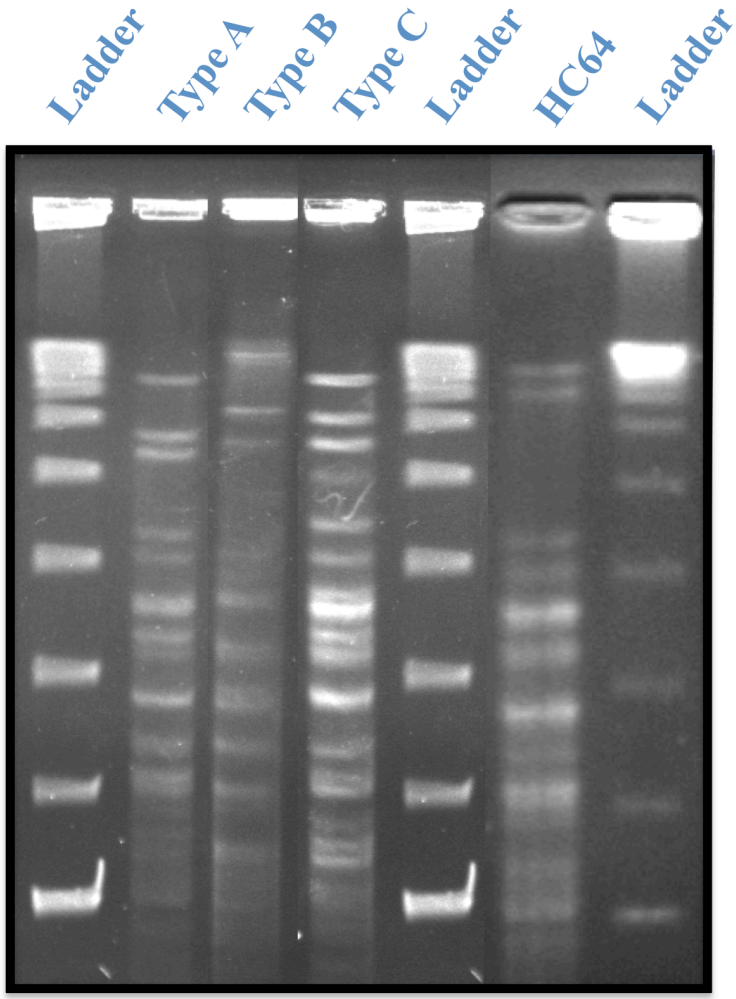


**Fig. S1 – Pulsed field gel of strains referenced in the manuscript.**

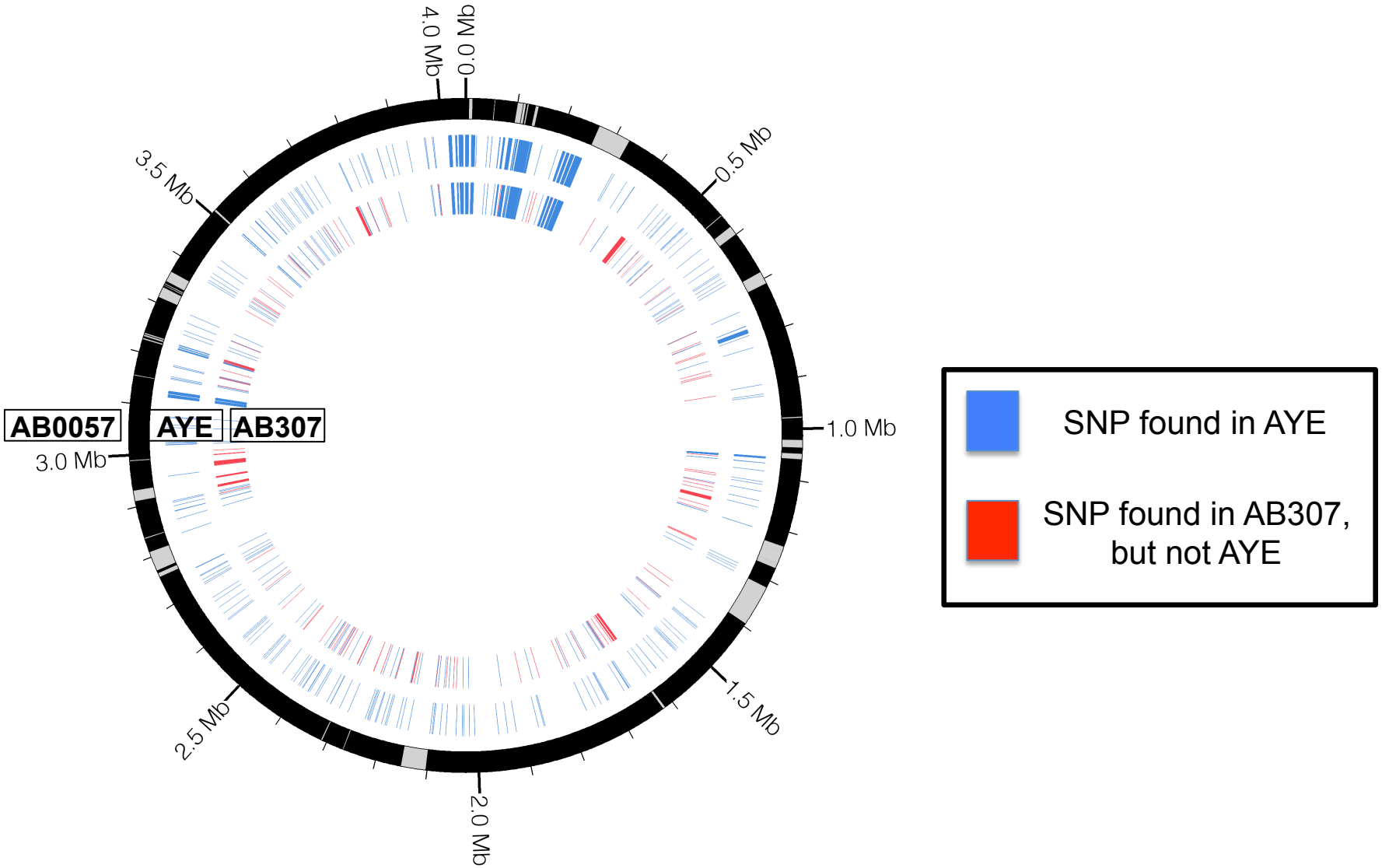


**Fig. S2 - Distribution of nucleotide differences among three sequenced genomes in the European clone I lineage.** Single nucleotide polymorphisms in the genome alignments of strains AYE and AB307-0294, relative to AB0057 were determined using Mauve. Positions differing from AB0057 (outer black circle) are indicated by colored marks in the two inner circles for AYE and AB307-0294, respectively. Different colors are used to represent shared alleles amongst different sets of genomes. Blue marks represent a position in which AYE differs from AB0057, with a blue mark also occurring in the circle for AB307-0294 if it shares this variant. Red marks represent a position in which AB307-0294 differs from both AYE and AB0057. In the outer circle, grey regions represent sequence present in AB0057, and not in at least one of the other two strains.

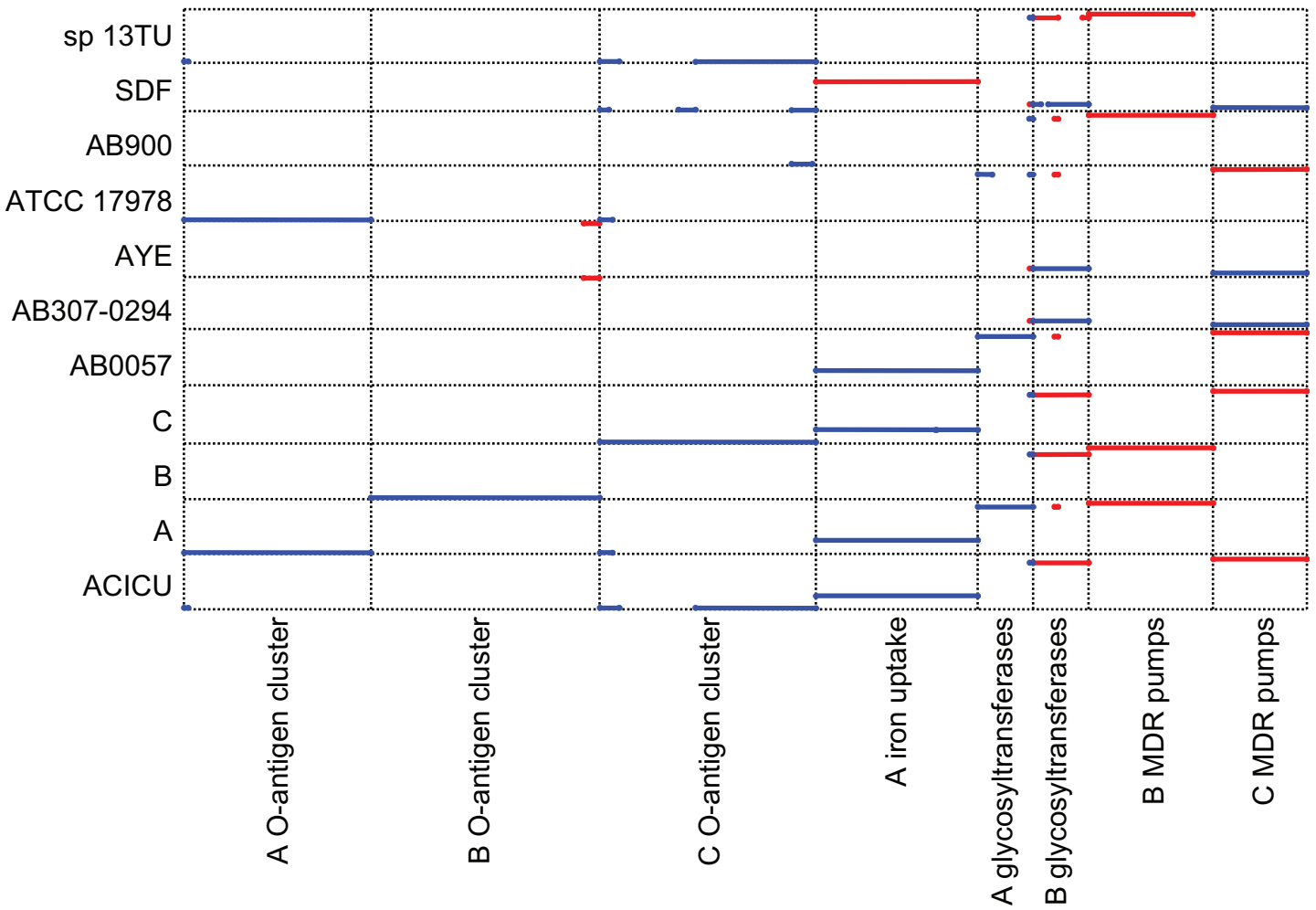


**Fig. S3 – Alignment of alternate recombinant regions against sequenced *A. baumannii* genomes.** The genomic sequences unique to recombinant regions that were discussed in the main text were manually extracted from the EC II genomes, and aligned against other *Acinetobacter* genomes using MUMmer (Supplementary reference 5). These alignments were displayed using the mummerplot function, associated with the MUMmer package. The columns correspond to the labeled recombinant regions, with the width of the column being proportional to the region size, and the rows similarly correspond to the *Acinetobacter* genomes to which the regions were aligned. When a region aligns to a given genome, this is indicated by the presence of a blue (forward) or red (reverse complement) line at the appropriate position in the matrix. The vertical position of the line within a given row is determined by the position in the genome to which the region aligns.

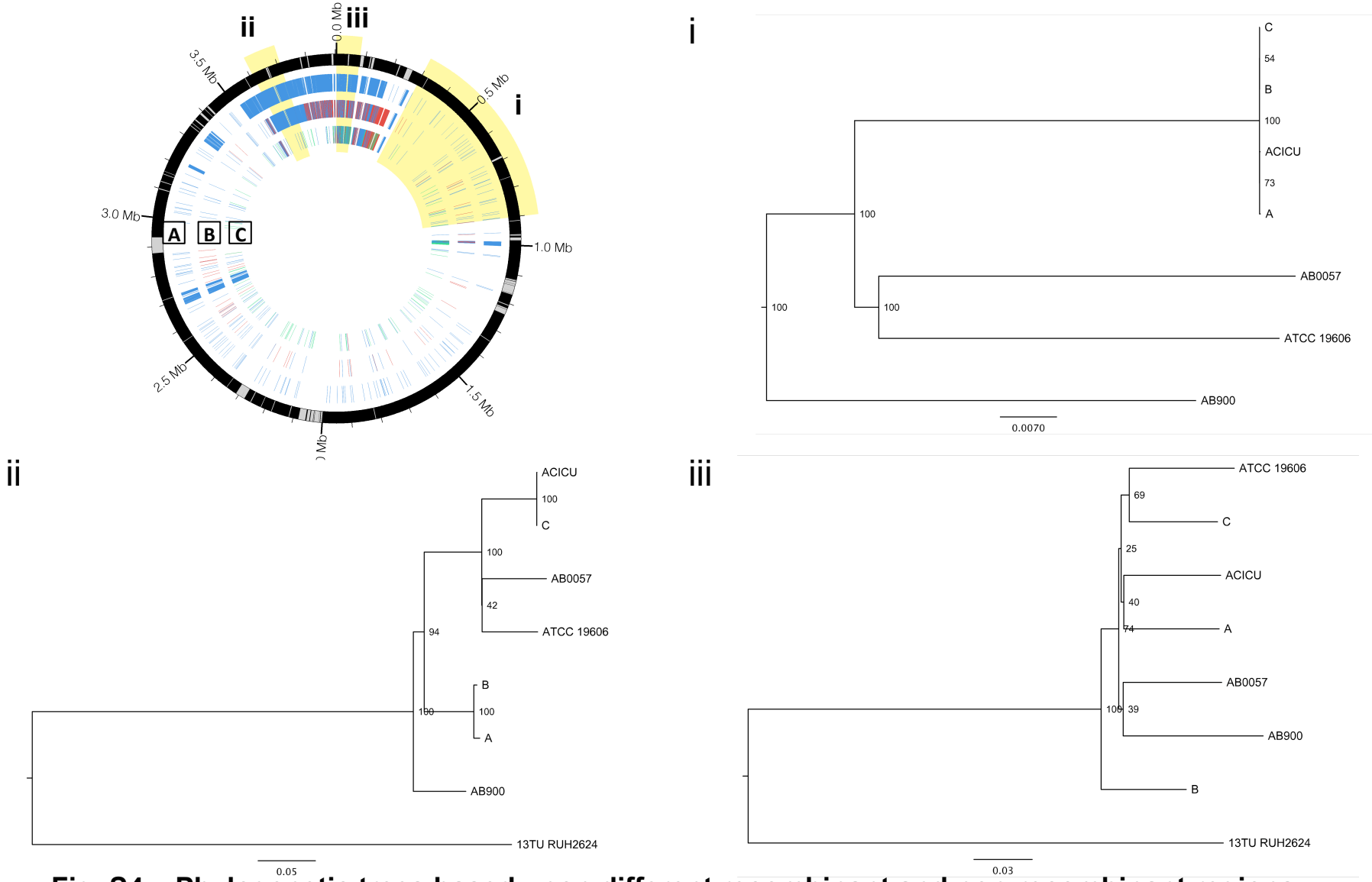


**Fig. S4 – Phylogenetic trees based upon different recombinant and non-recombinant regions.** The circle plot is the same as from Fig. 2 from the main text, and represents the distribution of SNPs among the outbreak genomes, relative to ACICU. Phylogenetic trees were constructed using the Neighbor Joining algorithm in the Seaview software package, based on SNPs calls from Mauve multiple genome alignments. The trees labeled i, ii and iii were constructed based upon SNPs in the highlighted regions of the circle plot.

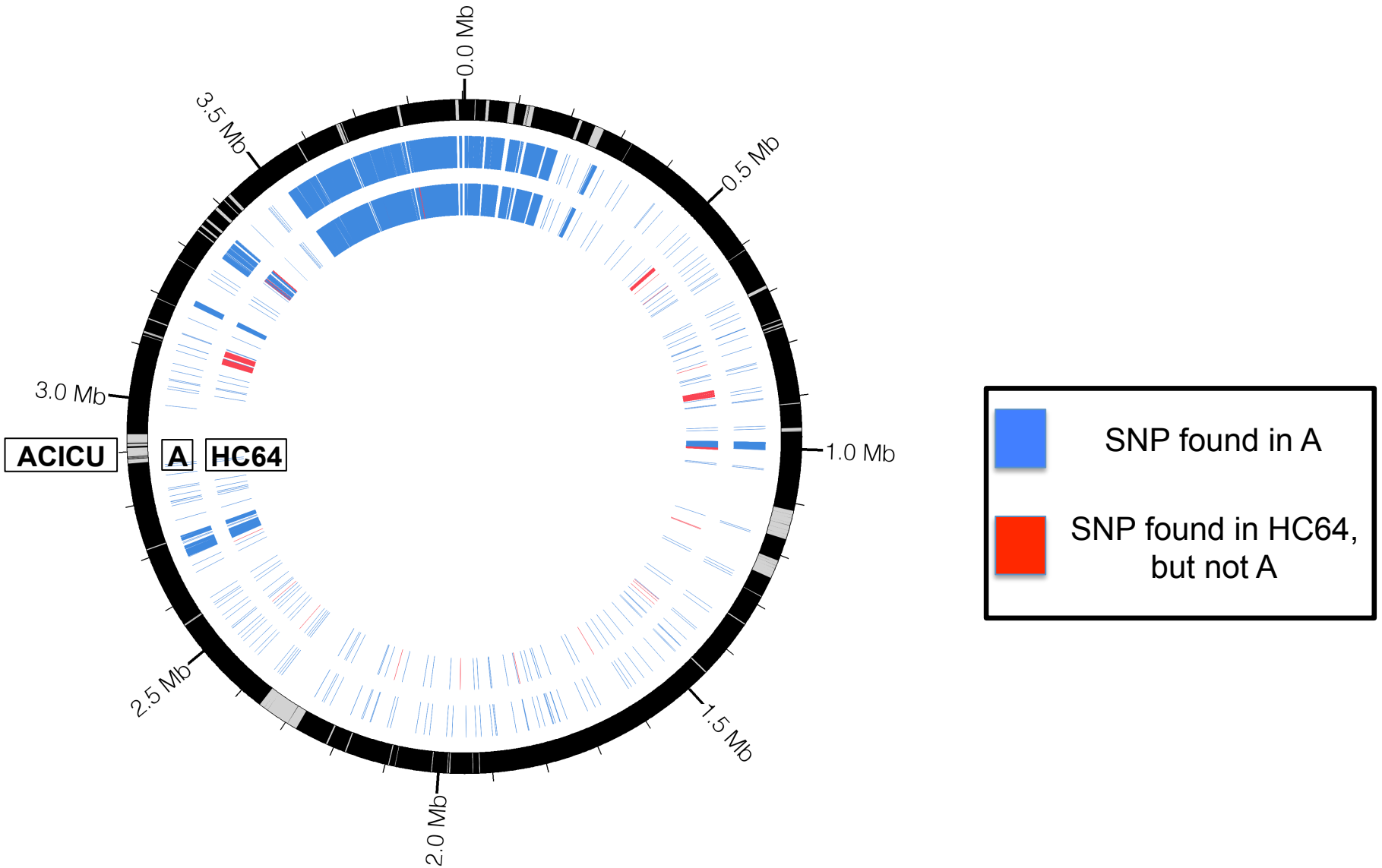


**Fig. S5 – Distribution of nucleotide differences between A, HC64 and ACICU.** Single nucleotide polymorphisms in the genome alignments of A and HC64, relative to ACICU were determined using Mauve. Positions differing from ACICU (outer black circle) are indicated by colored marks in the two inner circles for A and HC64, respectively. Different colors are used to represent shared alleles amongst different sets of genomes. Blue marks represent a position in which A differs from ACICU. If HC64 shares this variant allele with A, then a blue mark is displayed in its designated circle as well. Red marks represent a position in which HC64 differs from both A and ACICU. In the outer circle, grey regions represent sequence present in only ACICU.

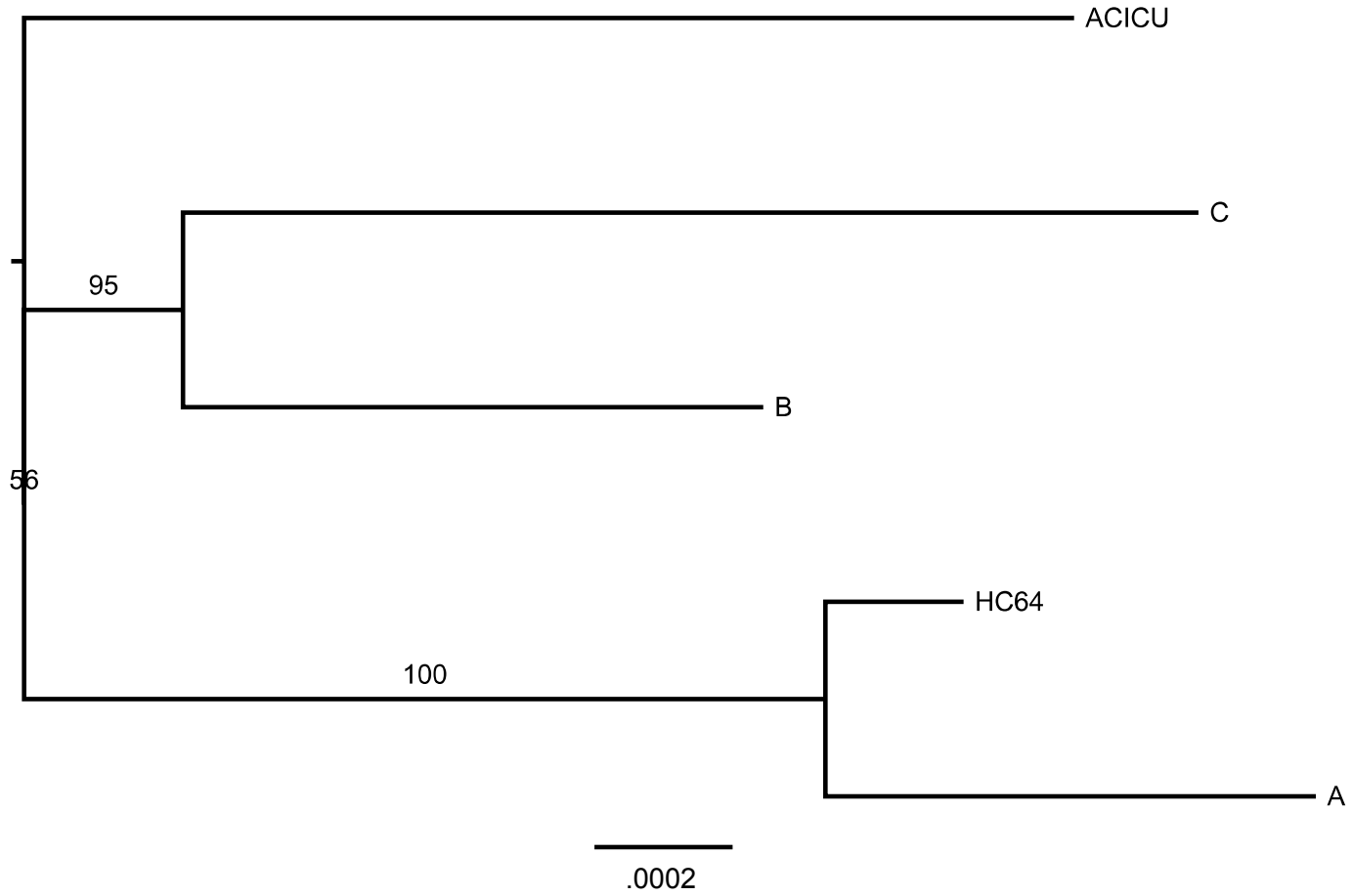


**Fig. S6- Phylogenetic relationship among EC II genomes based upon base differences outside of putative recombined regions.** The evolutionary relationship among strains A, B, C, ACICU and HC64 was determined based upon nucleotide differences that occur outside of putative recombined regions. A phylogenetic tree was constructed based upon these base differences using a Maximum Likelihood approach. Bootstrap values from 100 samplings are indicated on the appropriate nodes.

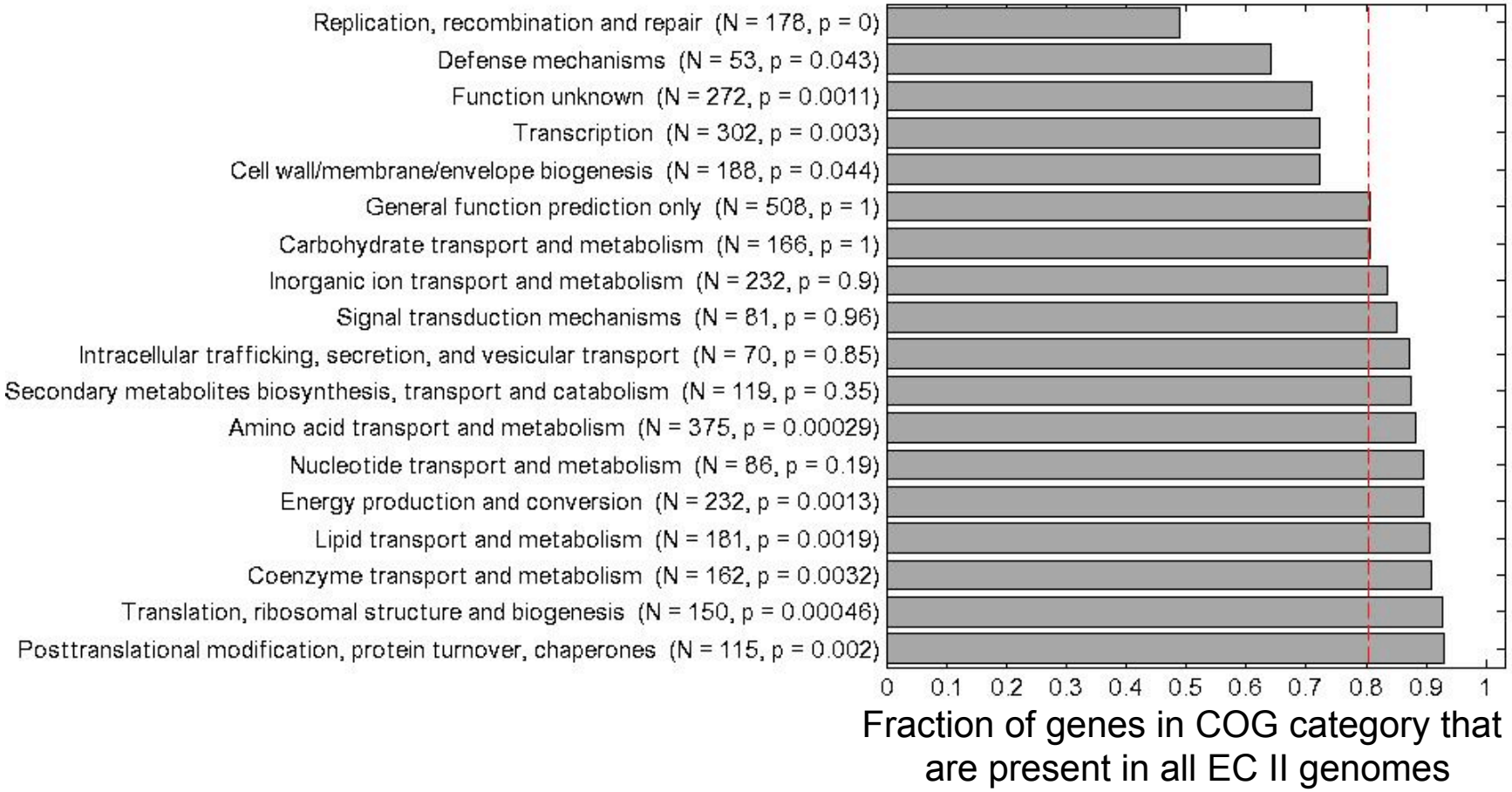


**Fig. S7- Functional enrichment among genes varying across four EC II genomes. (A)** Orthologous gene groups were determined for A, B, C and ACICU genomes. The fraction of genes in each COG functional category (y-axis) that are variable (i.e. not present in all four genome) is plotted on the x-axis. The dotted red line represents the null expectation for each functional category. To determine if any functional category is over- or under-represented, Fisher's exact tests were performed. The Bonferroni corrected two-sided p-values associated with the tests performed for each COG category are shown on the y-axis label.

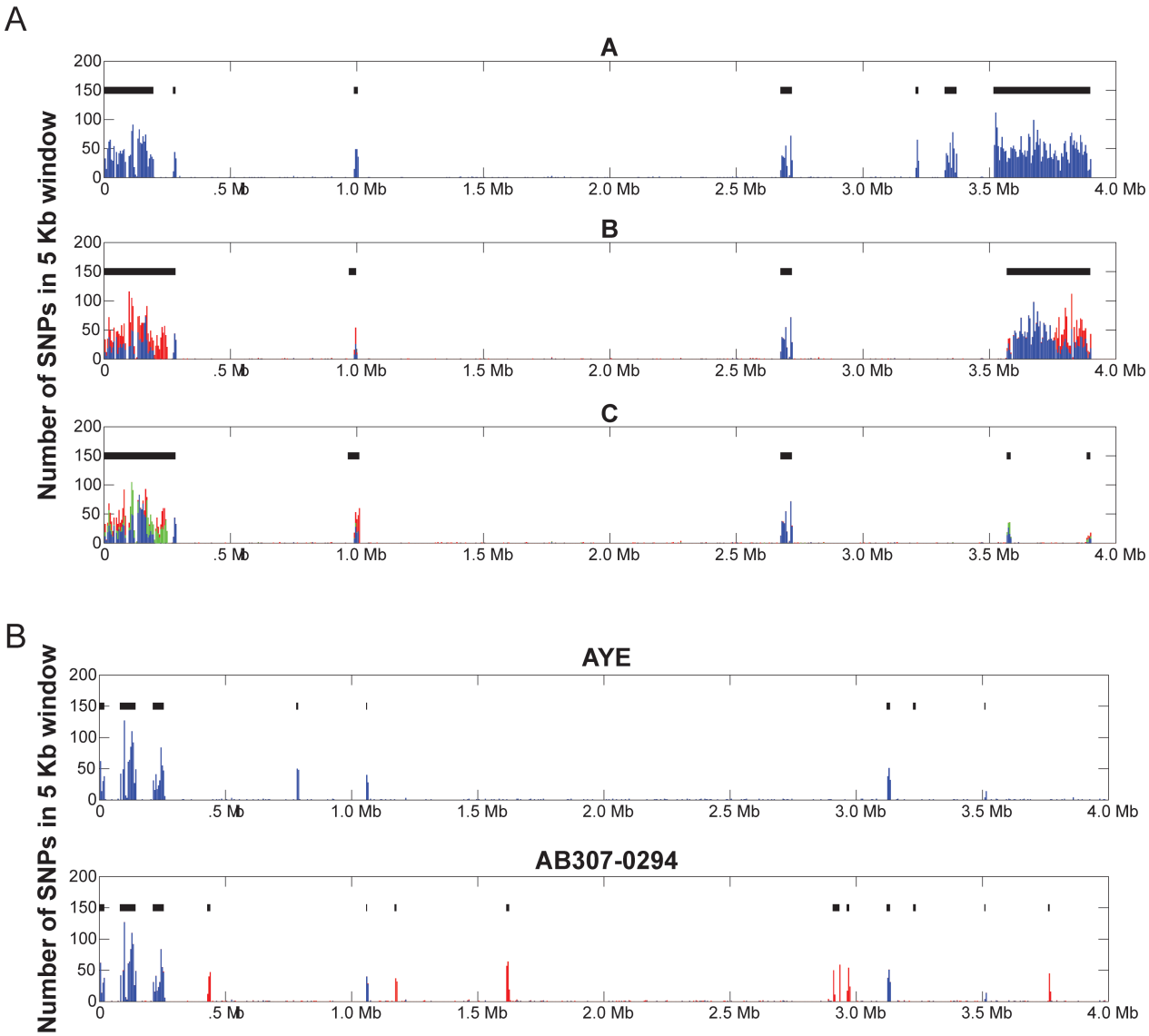


**Fig. S8- SNP density among EC I and EC II genomes in 5 Kb windows.** (A) Single nucleotide polymorphisms in the genome alignments of A, B and C relative to ACICU were determined using Mauve. The number of SNPs in A, B and C relative to ACICU were plotted as histograms in 5 Kb windows. Different colors represent shared alleles amongst different sets of genomes. Blue represents variants in which A differs from ACICU, with blue also occurring in the histograms for B and/or C if they also share these variants. Red represents variants in which B differs from both A and ACICU, with red also occurring in the histogram for C if it shares these variants. Finally, green represents variants in which C differs from ACICU, A and B. (B) SNPs were determined for AYE and AB307-0294 relative to AB0057 and plotted as histograms in 5 Kb windows. Again, different colors are used to represent shared alleles.

| Gene Set | Sequencing method | Gene | Fragment Length (bp) | # of differences between A and ACICU | # of differences between B and ACICU | # of differences between C and ACICU |
|---|---|---|---|---|---|---|
| http://www.pasteur.fr/mlst | genome | cpn60 | 405 | 0 | 0 | 0 |
| http://www.pasteur.fr/mlst | genome | fusA | 633 | 0 | 0 | 0 |
| http://www.pasteur.fr/mlst | genome | gltA | 433 | 0 | 0 | 0 |
| http://www.pasteur.fr/mlst | genome | pyrG | 297 | 0 | 0 | 0 |
| http://www.pasteur.fr/mlst | genome | recA | 372 | 0 | 0 | 1 |
| http://www.pasteur.fr/mlst | genome | rplB | 330 | 0 | 0 | 0 |
| http://www.pasteur.fr/mlst | genome | rpoB | 456 | 0 | 0 | 0 |
| http://pubmlst.org/abaumannii/ | MLST | gltA | 484 | 0 | 0 | 0 |
| http://pubmlst.org/abaumannii/ | MLST | gyrB | 457 | 4 | 3 | 2 |
| http://pubmlst.org/abaumannii/ | MLST | gdhB | 396 | 0 | 0 | 0 |
| http://pubmlst.org/abaumannii/ | MLST | recA | 371 | 0 | 0 | 1 |
| http://pubmlst.org/abaumannii/ | MLST | cpn60 | 421 | 0 | 0 | 0 |
| http://pubmlst.org/abaumannii/ | MLST | gpi | 305 | 6 | 9 | 6 |
| http://pubmlst.org/abaumannii/ | MLST | rpoD | 513 | 0 | 0 | 0 |

**Table S1 – Comparison of MLST alleles of outbreak strains to ACICU**

| Strain | Median/Mean Depth | Number of contigs | Contig N50 | Number of scafolds | Scaffold N50 | Length of concatenated scaffolds and contigs | Number of protein coding genes |
|---|---|---|---|---|---|---|---|
| A | 23/28.5 | 103 | 81821 | 20 | 385209 | 3953855 | 3697 |
| B | 19/21.8 | 125 | 78659 | 12 | 2564149 | 4014916 | 3809 |
| C | 11/16.5 | 284 | 35493 | 19 | 468134 | 4200364 | 4044 |
| HC64 | 17/18.1 | 184 | 43094 | NA | NA | 3996619 | 3891 |

**Table S2 – Summary of draft genomes statistics**

| | ACICU | A | B | C |
|---|---|---|---|---|
| **ACICU** | 0 | 0.28% | 0.26% | 0.13% |
| **A** | 9470 | 0 | 0.21% | 0.27% |
| **B** | 8736 | 7059 | 0 | 0.21% |
| **C** | 4466 | 9348 | 7224 | 0 |

**Table S3 – Number of nucleotide differences among European clone II lineage genomes**

| Region | Region size | Number of genes in region | Gene names | Annotation of best BLAST hit in non-redundant database | COG Annotations |
|---|---|---|---|---|---|
| A O-antigen cluster | 10935 | 10 | ABNIH1_07767 ,ABNIH1_07772 ,ABNIH1_07777 ,ABNIH1_07782 ,ABNIH1_07797 ,ABNIH1_07802 ,ABNIH1_07807 ,ABNIH1_07812 ,ABNIH1_07817 ,ABNIH1_07822 | WecC protein, MviM protein, WbbJ protein, glutamine--scyllo-inositol transaminase, capsular polysaccharide synthesis enzyme, glycosyltransferase, hypothetical protein, hypothetical protein, putative glycosyltransferase, hypothetical protein | UDP-N-acetyl-D-mannosaminuronate dehydrogenase ,Predicted dehydrogenases and related proteins ,Acetyltransferase (isoleucine patch superfamily) ,Predicted pyridoxal phosphate-dependent enzyme apparently involved in regulation of cell wall biogenesis ,Acetyltransferase (isoleucine patch superfamily) ,Glycosyltransferase ,Glycosyltransferase , ,Glycosyltransferase ,Glycosyltransferases involved in cell wall biogenesis |
| B O-antigen cluster | 13338 | 11 | ABNIH2_18345 ,ABNIH2_18350 ,ABNIH2_18355 ,ABNIH2_18360 ,ABNIH2_18365 ,ABNIH2_18370 ,ABNIH2_18375 ,ABNIH2_18380 ,ABNIH2_18385 ,ABNIH2_18390 ,ABNIH2_18395 | polyprenol phosphate:N-acetyl-hexosamine 1-phosphate transferase, UDP-glucose 4-epimerase, glycosyl transferases group 1-like protein, UDP-N-acetylglucosamine 2-epimerase, WbjC, WbjB, group 1 glycosyl transferase, glycosyl transferase group 1, hypothetical protein, O-antigen translocase, Vi polysaccharide biosynthesis protein vipB/tviC | UDP-N-acetylmuramyl pentapeptide phosphotransferase/ UDP-N-acetylglucosamine-1-phosphate transferase ,Nucleoside-diphosphate-sugar epimerases ,Glycosyltransferase ,UDP-N-acetylglucosamine 2-epimerase ,Nucleoside-diphosphate-sugar epimerases ,Predicted nucleoside-diphosphate sugar epimerases ,Glycosyltransferase ,Glycosyltransferase , ,Membrane protein involved in the export of O-antigen and teichoic acid ,Nucleoside-diphosphate-sugar epimerases |
| C O-antigen cluster | 12636 | 13 | ABNIH3_04449 ,ABNIH3_04454 ,ABNIH3_04459 ,ABNIH3_04464 ,ABNIH3_04469 ,ABNIH3_04474 ,ABNIH3_04479 ,ABNIH3_04484 ,ABNIH3_04489 ,ABNIH3_04494 ,ABNIH3_04499 ,ABNIH3_04504 ,ABNIH3_04509 | UDP-N-acetyl-D-mannosaminuronate dehydrogenase, nucleoside-diphosphate sugar epimerase, pyridoxal phosphate-dependent enzyme, CMP-N-acetylneuraminic acid synthetase, spore coat polysaccharide biosynthesis protein, glycosyltransferase, acetyltransferase, sialic acid synthase, polysaccharide biosynthesis protein, hypothetical protein, polysaccharide biosynthesis protein, hypothetical protein, hypothetical protein, glycosyltransferase | UDP-N-acetyl-D-mannosaminuronate dehydrogenase ,Predicted nucleoside-diphosphate sugar epimerases ,Predicted pyridoxal phosphate-dependent enzyme apparently involved in regulation of cell wall biogenesis ,CMP-N-acetylneuraminic acid synthetase ,Spore coat polysaccharide biosynthesis protein, predicted glycosyltransferase ,Acetyltransferases, including N-acetylases of ribosomal proteins ,Sialic acid synthase ,Membrane protein involved in the export of O-antigen and teichoic acid , , , ,Zn-dependent hydrolases, including glyoxylases ,Glycosyltransferases involved in cell wall biogenesis |
| A iron uptake genes | 9450 | 7 | ABNIH1_03437 ,ABNIH1_03442 ,ABNIH1_03447 ,ABNIH1_03452 ,ABNIH1_03457 ,ABNIH1_03462 ,ABNIH1_03467 | DNA-directed RNA polymerase specialized sigma subunit, putative transmembrane sensor protein FecR, receptor protein; Heme acquisition system receptor, signal peptide, hypothetical protein, putative TonB-related protein, heme oxygenase | DNA-directed RNA polymerase specialized sigma subunit, sigma24 homolog ,Fe2+-dicitrate sensor, membrane component ,Outer membrane receptor proteins, mostly Fe transport , ,FOG: TPR repeat ,Periplasmic protein TonB, links inner and outer membranes ,Heme oxygenase |
| A glycosyltransferases | 3240 | 3 | ABNIH1_00190 ,ABNIH1_00205 ,ABNIH1_00210 | LPS glycosyltransferase subfamily protein, glycosyltransferase, glycosyltransferase | ,Glycosyltransferase , |
| B glycosyltransferases | 3240 | 4 | ABNIH2_11953 ,ABNIH2_11958 ,ABNIH2_11963 ,ABNIH2_11968 | glycosyltransferase, hypothetical protein, hypothetical protein, putative glycosyltransferase | , , ,Glycosyltransferase involved in LPS biosynthesis |
| B MDR pumps | 7263 | 5 | ABNIH2_14445 ,ABNIH2_14450 ,ABNIH2_14455 ,ABNIH2_14460 ,ABNIH2_14465 | membrane-fusion protein, outer membrane protein precursor, component of the czc cation-efflux system (CzcI), hypothetical protein, nickel-cobalt-cadmium resistance protein | Membrane-fusion protein ,Outer membrane protein , , ,Putative protein-S-isoprenylcysteine methyltransferase |
| C MDR pumps | 5481 | 2 | ABNIH3_19200 ,ABNIH3_19205 | membrane-fusion protein, outer membrane protein | Membrane-fusion protein ,Outer membrane protein |

**Table S4 – Gene content in recombinant regions associated with gene turnover in the EC II lineage**

| Strain Code | Date | MDR | Pulsotype | LPS A | LPS B | LPS C |
|---|---|---|---|---|---|---|
| AC87 | 5/24/07 | + | A | YES | | |
| *AC00* | *7/4/07* | + | A | YES | | |
| *BD59* | *9/15/07* | + | B | | YES | |
| BD30 | 11/1/07 | + | B | | YES | |
| *AR82* | *8/28/07* | + | C | | | YES |
| *AR21* | *1/30/08* | + | NT | | YES | |
| *HB77* | *7/12/07* | + | B | | YES | |
| *HB57* | *7/12/07* | + | C | | | YES |
| *BJ88* | *7/29/07* | + | B | | YES | |
| *BJ38* | *8/17/07* | + | C | | | YES |
| *TD11* | 8/25/07 | + | A | YES | | |
| *TD72* | *9/3/07* | + | B | | YES | |
| *EB68* | *7/23/07* | + | A | YES | | |
| *EB51* | *12/19/08* | + | B | | YES | |
| | | | | | | |
| *ML38* | *11/9/06* | - | Unique | | | |
| *WJ40* | *5/25/06* | - | Unique | | | |
| *HC64* | *2/14/06* | - | Unique | YES | | |
| *DC42* | *5/5/06* | - | Unique | | | |
| *GW41* | *2/16/06* | - | Unique | | | |

**Table S5 – Results of PCR assay to assess the identity of the O-antigen cluster among isolates taken at the NIHCC in the time before and after the outbreak began**

| **Drug** | **A** | **B** | **C** | **HC64** |
|---|---|---|---|---|
| Amikacin | <=16 (S) | >32 (R) | >32 (R) | <=16 (S) |
| Amp/Sulbactam | NA | >16/8 (R) | >16/8 (R) | NA |
| Aztreonam | >16 (R) | >16 (R) | >16 (R) | >16 (R) |
| Cefepime | 16 (I) | >16 (R) | 16 (I) | >16 (R) |
| Ceftazidime | >16 (R) | >16 (R) | >16 (R) | >16 (R) |
| Ciprofloxacin | >2 (R) | >2 (R) | >2 (R) | >2 (R) |
| Colistin | 0.25 (S) | 0.5 (S) | 0.125 (S) | NA |
| Gentamicin | >8 (R) | >8 (R) | >8 (R) | <=4 (S) |
| Imipenem | >32 (R) | >8 (R) | <=4 (S) | <=4 (S) |
| Levofloxacin | >4 (R) | >4 (R) | >4 (R) | >4 (R) |
| Meropenem | >8 (R) | >8 (R) | 8 (I) | 8 (I) |
| Pip/Tazo | >64 (R) | >64 (R) | >64 (R) | >64 (R) |
| Rifampin | NA | 4 | 4 | NA |
| Ticar/K Clavulanate | 64 (I) | >64 (R) | >64 (R) | 64 (I) |
| Tigecycline | 4 | 2 | 2 | NA |
| Tobramycin | <=4 (S) | >8 (R) | >8 (R) | <=4 (S) |
| Trimeth/Sulfa | >2/38 (R) | >2/38 (R) | >2/38 (R) | <=2/38 (S) |

**Table S6 – Antibiotic resistance profiles of sequenced strains**

| O-antigen locus | Forward primer | Reverse primer | Amplicon size (bp) | Anneal temperature (C) | Extension time (min) | Number of cycles |
|---|---|---|---|---|---|---|
| A | TGGCTACGAATGGGAAAAGGTCGG | ACGACTGCTCCAGCAGCAACT | 872 | 55 | 1:30 | 28 |
| B | GTCGACGATGTAGTTGAAACATTCT | TATCTTTGCCAATGACAATGACCTC | 2250 | 55 | 2:30 | 30 |
| C | AAGTAGAGCTGAAAATAGTAGTTGA | TCAGACCATACTTTAGATAATACTAC | 2026 | 55 | 2:15 | 30 |

**Table S7 – PCR primers and conditions for probing the presence of O-antigen cluster of A, B and C**